Metáfora de Bill Gates en un bar

Nos permite diferenciar los estadísticos.

Formulación matemática

$$dataset = \{x_1, x_2, ...X_n\}$$

$$media = \overline{X} = \frac{\{x_1, x_2, ...X_n\}}{N} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

Hay 2 casos para obtener la mediana, ya que puede haber un número $par\ o\ impar$ de números

Tenemos que:

- Par

$$x = \{1, 2, 3\} \rightarrow length(x) = 3 = N$$

- Impar

$$x = \{1, 2, 3, 4\} \rightarrow length(x) = 4 = N$$

Tenemos que tener ordenado el conjunto y podemos obtener la mediana de la siguiente forma

$$mediana \quad \begin{cases} \text{Par} & \frac{X_{\left[\frac{N}{2}\right]} + X_{\left[\frac{N}{2}+1\right]}}{2} \\ \text{Impar} & X_{\left[\frac{N+1}{2}\right]} \end{cases}$$

Metáfora de Bill Gates

- Hay 11 personas en un bar que trabajan para una empresa, cada una tiene un salario de $35,000.Calculemoslos2$\$ parámetros anteriores:

$$media = \overline{X} = \frac{(11)(\$35,000)}{11} = \$35,000$$

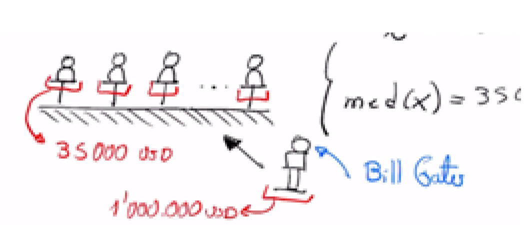
$$mediana = \frac{\$35,000 + \$35,000}{2} = \$35,000$$

Entonces de acuerdo a que todos ganan lo mismo, los 2 parámetros tienen el mismo valor.

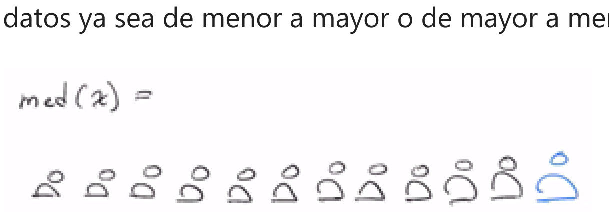
- Ahora llega Bill Gates al bar y se sienta junto a los trabajadores. Pero el no tiene el mismo salario, el gana \$1,000,000. Alguien propone calcular los mismos parámetros y se procede a realizar las operaciones para saber cuál es el salario promedio entre las personas del bar.

$$\overline{X} = \frac{(11)(\$35,000) + \$1,000,000}{12} = \$115,000$$

Esto nos dice que $\$115,000$ es el salario promedio entre las personas que hay en el bar, pero digamos el hecho te alguien tenga un salario elevado hace que parezca que las personas restantes ganan mucho más de lo que realmente están ganando.



Bill Gates es el muñequito en azul, ahora procederemos a calcular la $mediana$ veamos que pasa. Tengo que tener ordenados mis datos ya sea de menor a mayor o de mayor a menor.



Debido a que tenemos $n = 12$ es decir el numero de datos es par, tendremos que usar la formula para el caso de numero par:

$$mediana(x) = \frac{X_{\left[\frac{N}{2}\right]} + X_{\left[\frac{N}{2}+1\right]}}{2}$$

$$mediana(x) = \frac{\$35,000 + \$35,000}{2} = \$35,000$$

Esto está bien y la conclusion es que la mediana del salario es $35,000.

NOTA

En un conjunto donde hay un valor atípico, en este caso el salario de Bill Gates, hacia que el salario de las personas restantes pareciera elevado y esto puede dar un impresión equivocada, cuando en realidad las personas ganan menos de lo que decía el valor $promedio$ del análisis.

Entonces el promedio se ve *sesgado o desviado* por valores atípicos, entonces para este caso la $mediana$ es mejor para el manejo de esos valores atípicos en un conjunto de datos o en una distribución de datos.

Entonces pasa lo mismo con el Ingreso Per Capita de un país, ya que pueden existir casos en donde hayan valores atípicos como en el ejemplo, lo que en realidad no te mostraría la realidad de la situación.

Algunas medidas

Nombre	Fórmula
Media	$\bar{x} = \frac{1}{N} \cdot \sum_{i=1}^{N} x_i$
Moda	Datos x_i más repetidos.
Mediana	$x_{\frac{N+1}{2}}$ si N impar $\frac{1}{2} \cdot \left(x_{\frac{N}{2}} + x_{\frac{N}{2}+1} \right)$ si N par
Desviación respecto de la media	$D_i = \lvert x_i - \bar{x} \rvert$
Desviación media	$D_m = \frac{1}{N} \cdot \sum_{i=1}^{N} \lvert x_i - \bar{x} \rvert$
Varianza	$\sigma^2 = \frac{1}{N} \cdot \sum_{i=1}^{N} (x_i - \bar{x})^2$
Desviación estándar	$\sigma = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} (x_i - \bar{x})^2}$